AMD

HPC Day

AdamsMark November 4, 2003

Agenda

8:30	Cont. breakfast	
9:00	AMD Welcome	Tom Toles, AMD
9:15	Overview of AMD64	Douglas O'Flaherty, AMD
9:45	PGI Compiler Optimization	Doug Miles, PGI
10:45	Coffee Break	
11:00	Optimization with the ACML	Tim Wilkens, PhD, AMD
12:00 – 1:00pm	Lunch	
1:00	The new Red Hat releases	Dave Johnson, Red Hat
1:30	Coding with OpenGL	Cass Everitt, Nvidia
2:30	Parallel Debugging	Doug Miles, PGI

AMD64

Douglas O'Flaherty

HPC Day November 4, 2003

AMD Opteron™ Processor Wins !

United Defense
6 Systems

UCIrvine University of California, Irvine
20 Systems

曙光 DAWNING China
>500 Systems

U.S. AIR FORCE
16 Systems

20 Systems

Stanford University
24 Systems

DaimlerChrysler
152 Systems

AIST
1,058 Systems

Virginia Tech
200 Systems

weta Ltd. visual effects

Los Alamos NATIONAL LABORATORY
256 Systems

AMERICA'S ARMY
88 Systems

256 Systems

Los Alamos NATIONAL LABORATORY
1,250 Systems

NCSOFT
20+ Systems

BROOKHAVEN NATIONAL LABORATORY
2 Systems

20 Systems

THE UNIVERSITY of MANCHESTER
35 Systems

CRAY

Sandia National Laboratories
10,136 Systems

ROSE-HULMAN INSTITUTE OF TECHNOLOGY
2 Systems

NASA JPL
152 Systems

Kasetsart University
20 Systems / QTR

Recent AMD Opteron™ Processor Wins

Sandia National Laboratories

CRAY

- ✓ **Cray plans to build a 40+ teraflop super computer using x86-64 AMD Opteron™ processors for Sandia National Laboratories**
- ✓ **Will be used for advanced engineering simulations**
- ✓ **$90 million project will use more than 10,000 AMD Opteron™ processors**
- ✓ **Design features a simple building block approach with HyperTransport™ technology that will enable easy implementation and reduce engineering, design, and component costs**

Recent AMD Opteron™ Processor Wins

- Lightening will be part of ASCI and will use 2,800 processors
  - To support the National Nuclear Security Administration's ASCI (Advanced Simulation and Computing) program
  - Approximately 11.2 TF peak performance
- Orange will be part of Institutional Computing
  256 nodes 512 AMD Opteron processor 244
  - To research the design of antibiotics and simulations of wildfires
  - Will use InfiniBand interconnect
  - Approximately 2 TF peak performance
- Announced August 14, 2003
- Vendor is Linux Networx

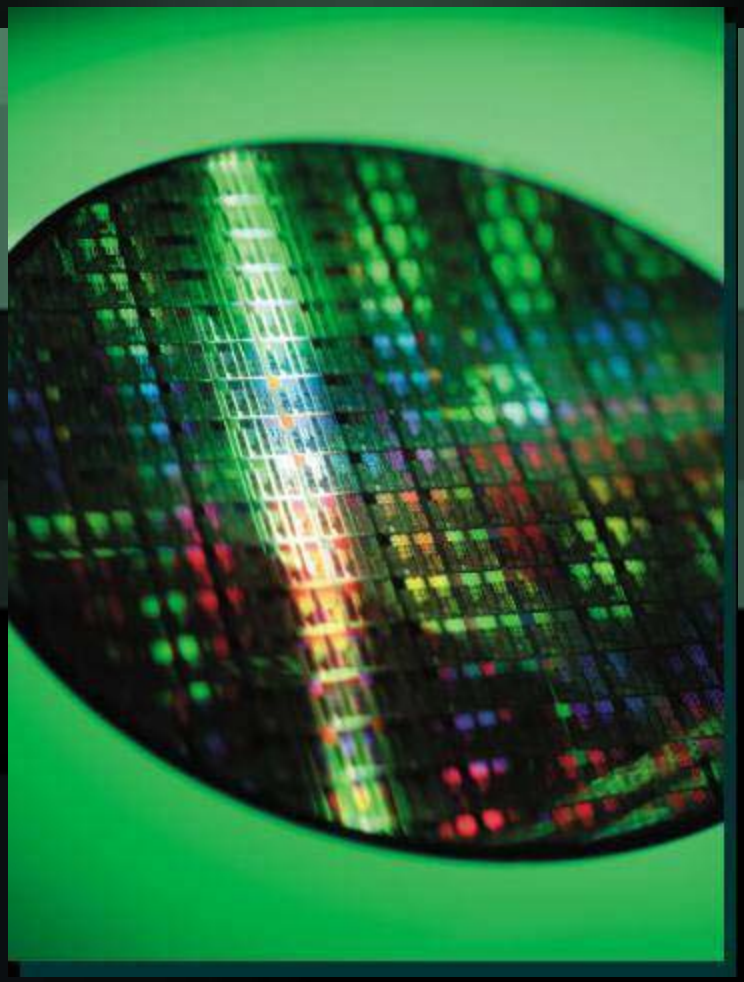
Recent AMD Opteron™ Processor Wins

- Japan's AIST
  - National Institute of Advanced Industrial Science and Technology (AIST) is Japan's largest national research organization
  - AIST will use 1,058 IBM eServer 325 systems powered by 2,116 AMD Opteron processor 246
    - To accelerate research using Grid computing in:
      - the search for new materials to be used for superconductors and fuel cell batteries, and
      - the search for new cures for various malignant diseases
      - Will use Globus Toolkit 3.0 and the Open Grid Services Infrastructure
  - Approximately 8.5 TF peak performance
  - Announced July 30, 2003
  - Vendor is IBM

Recent AMD Opteron™ Processor Wins

- Dawning 4000A

  - More than 2,000 AMD Opteron 800-series processors are planned to power the system with planned memory capacity up to 2.256 Terabytes. [4 processor node size]
  - The AMD Opteron processor is the only processor in the world capable of supporting a large, centrally located 64-bit computing resource while being compatible with existing 32-bit applications and the commodity desktop PC infrastructure.
  - End users for supercomputing in China include scientific research institutes, commercial establishments, private enterprises, and computing centers.
  - Expected to be greater than 10 TF peak performance
  - Announced July 28, 2003



AMD64
Design Choices

Review: Scale-Up

One system, [mid-range is typically SMP configuration]:

- Multiple processors (4P to 16P+)
- Large amounts of shared memory
- Shared I/O resources
- High bandwidth and low latency system interconnect
- OS must handle the balancing of incremental resources causing the complexity of OS to scale with processors and memory

- Fits a number of legacy commercial application scenarios where close sharing of memory by multiple worker threads is critical.
- 2P system that incrementally scales to 16P can cost 5x that of a 2P system that scales to 4P.
  - You pay now for the ability to scale-up later.
  - Systems that scale-up are more costly due to the complexity of the system, its interconnect, and RAS features.

Review: Scale-Out

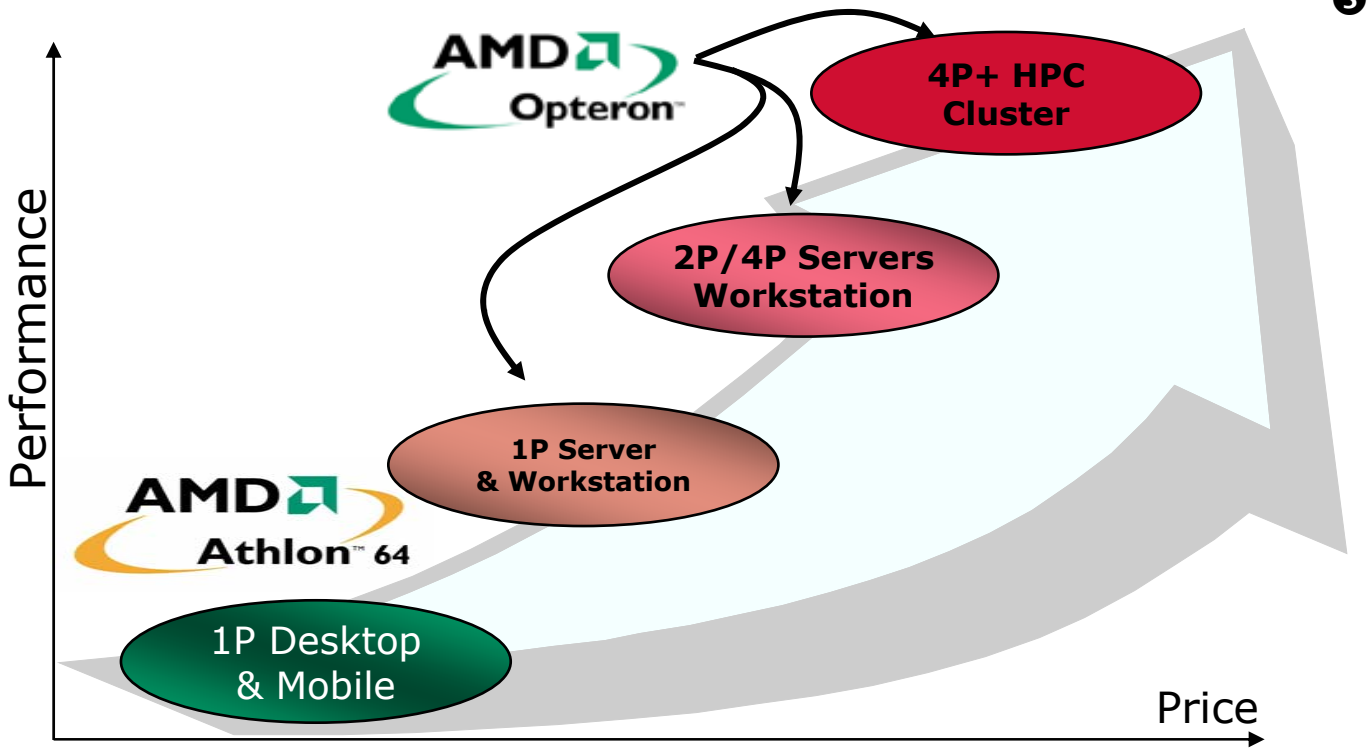
Many small, simple systems:

- Each system has 2 to 4 processors
  - Each system has its own memory
    - Systems are interconnected by a common network

- Each individual system does not need to be expensive, complex, and have a high-tolerance design
- Small size allows modular and dense packing in racks – very flexible
- Many RAS requirements can be satisfied by simple redundancy and ease of replacing/upgrading individual
- Network parallelism much easier for standard OS to handle
- However, the network is relatively slow interconnect:
  - Good fit for applications/workloads which can be decomposed into multiple threads/tasks with little data-sharing or communication.
  - Fits scientific/technical computing well

Targeting the Benefits of Both

AMD is targeting the mid-range, high-volume server space by supporting the best of both scale-up and scale-out approaches

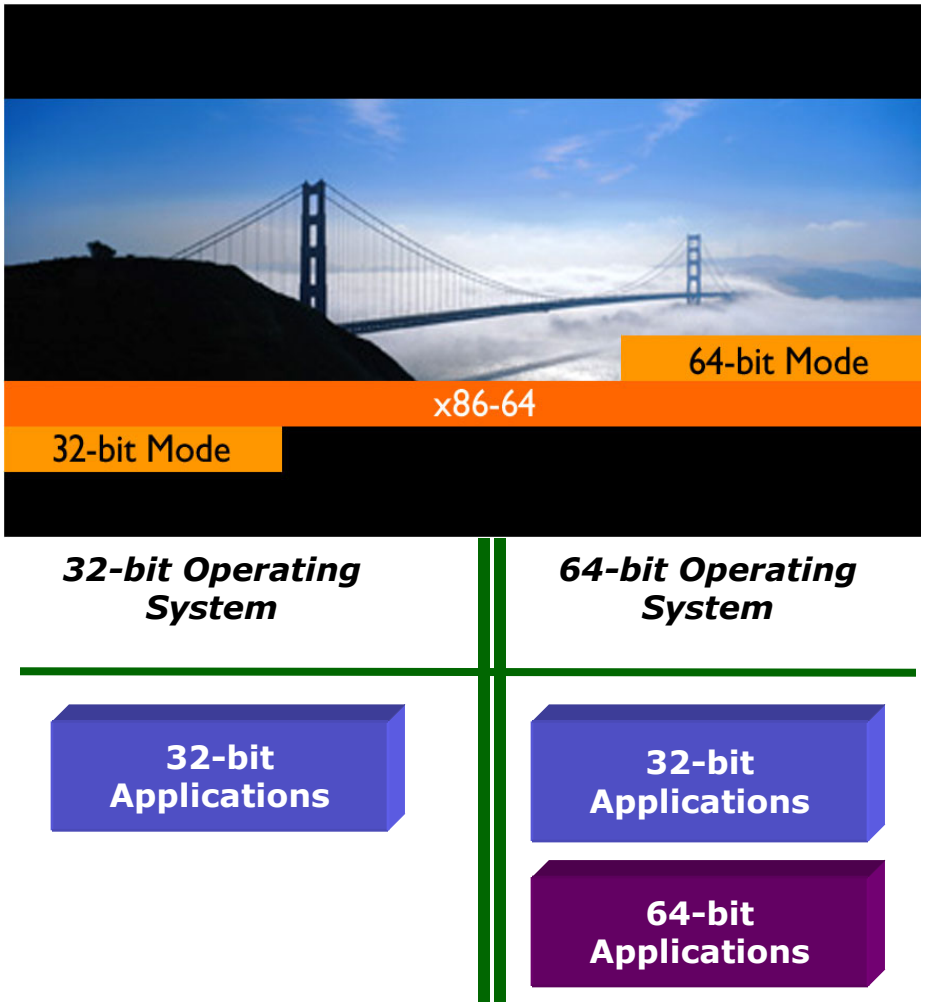


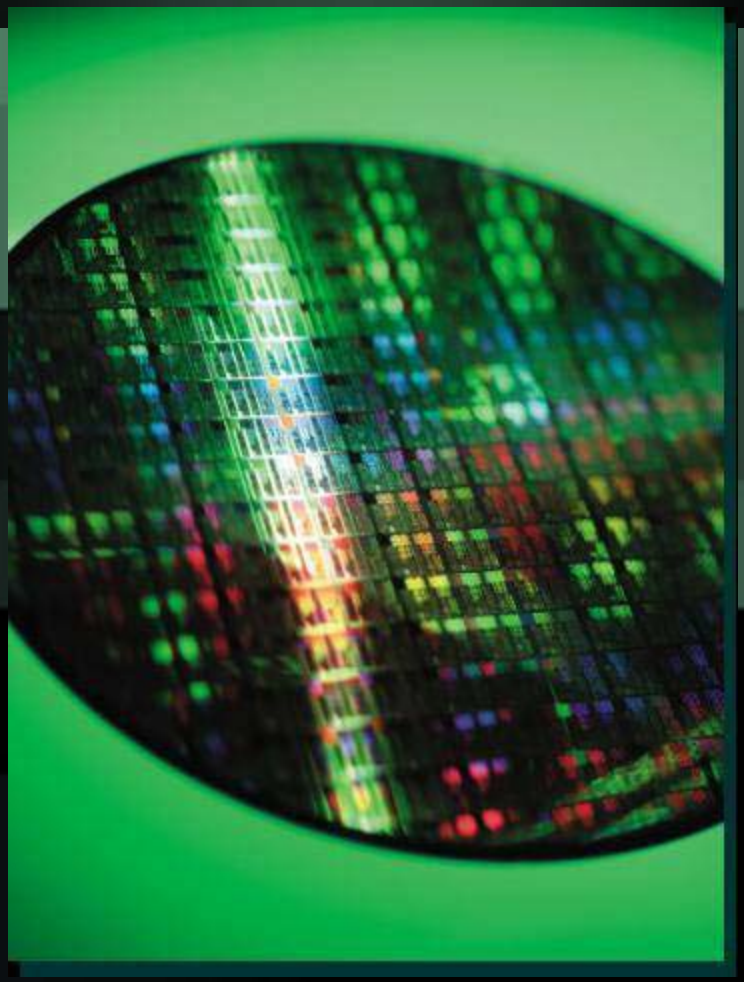
- ❺ Leverage & extend the Industry investment and expertise in x86

AMD64 Technology

Building a Bridge from the 32- to the 64-bit World

- Leverages the initial success of AMD Athlon™ MP processor
- Adds 64-bit capabilities to the world's highest performing 32-bit core for 2P and 4P servers
- Developed so that current 32-bit applications will work on both 32-bit and 64-bit operating systems
- Doesn't require special hardware or investment in a proprietary infrastructure
- Developing a solid ecosystem of motherboards, operating systems, development tools, and device drivers





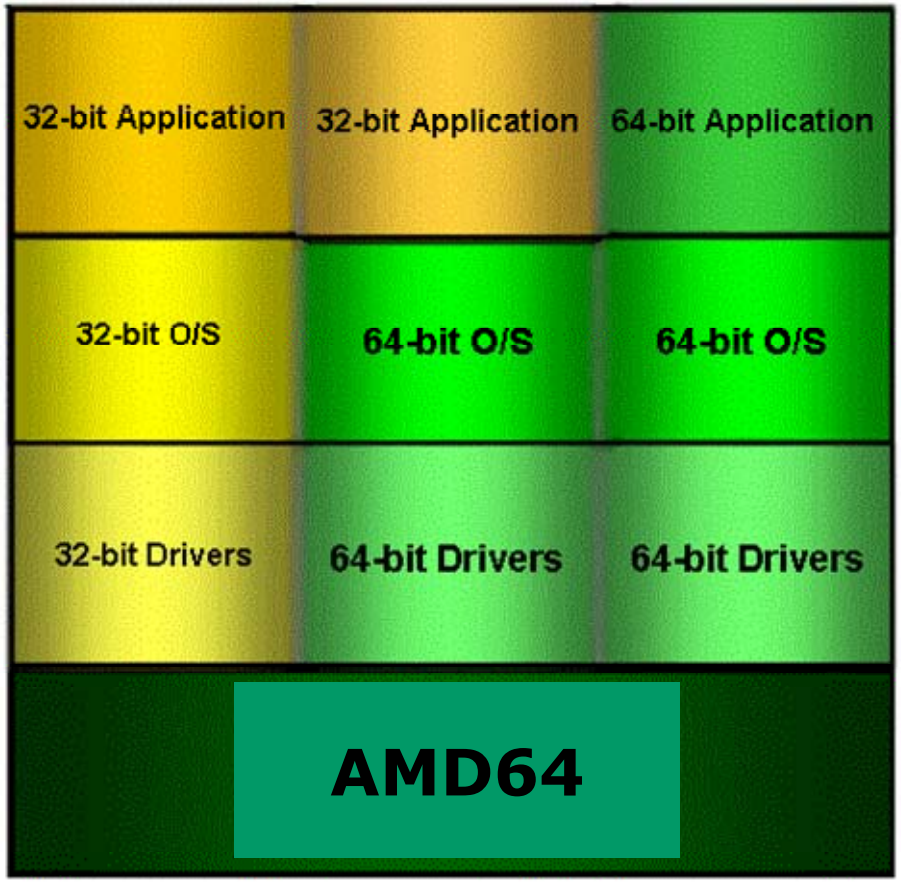
AMD64 Architecture

AMD64 Computing Strategy

AMD took the x86 architecture and extended it to 64-bits to make the AMD64 architecture

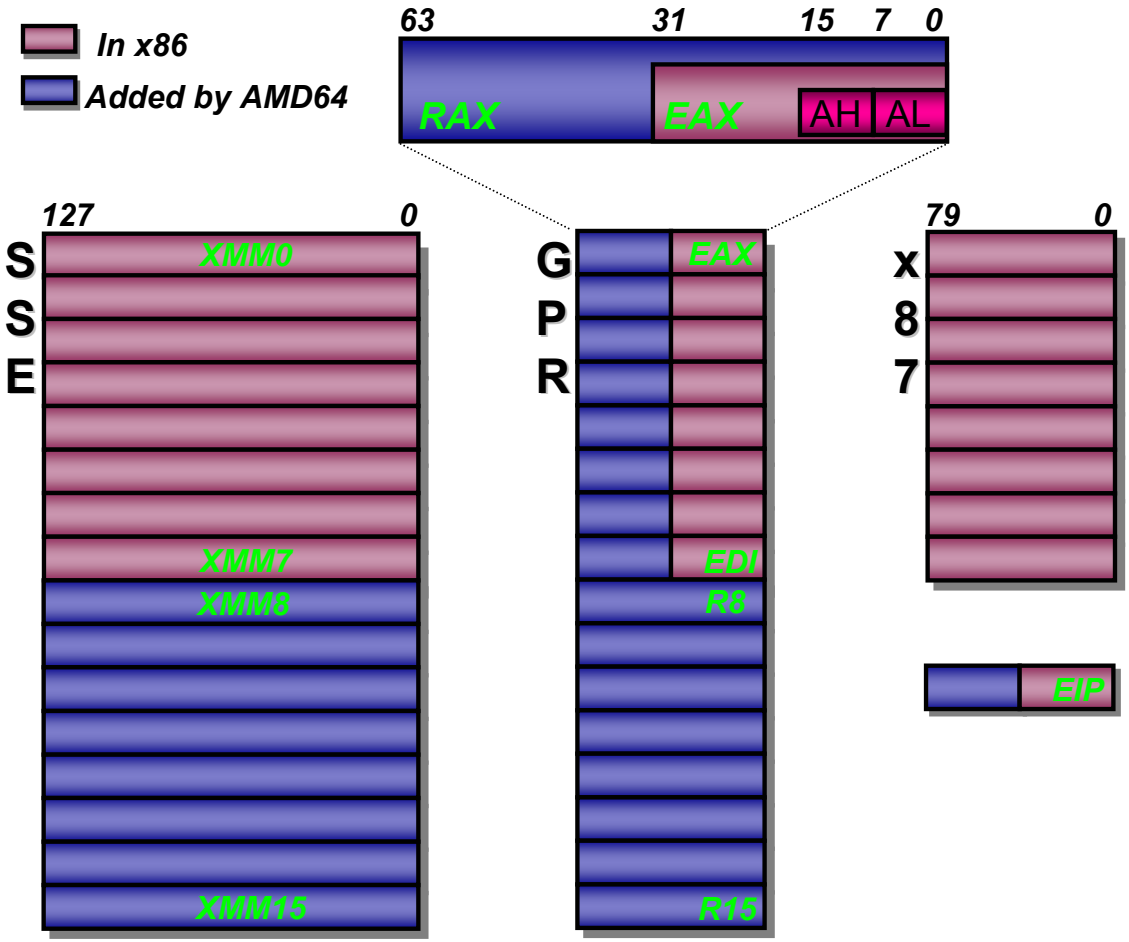
Extensions are so simple and compatible, that the processor can support both x86-32 and x86-64 at full speed & performance

- Offers native support and performance for 32-bit applications (legacy mode)
- Can move to 64-bit addressing and data types without giving up 32-bit compatibility (long mode)
  - Leverages the established infrastructure rather than needing to re-invent it.



AMD64 Computing Strategy

- AMD64 Architecture:
  - 64-bit integer registers
  - 64-bit Virtual Address
  - 52-bit Physical Address
  - Sixteen 64-bit integer regs
  - Sixteen 128-bit SSE regs
  - SSE2 Instruction Set
  - Double precision scalar and vector operations
  - 16x8, 8x16 way vector packed integer operations
  - SSE1 already added with AMD Athlon™ MP Processor



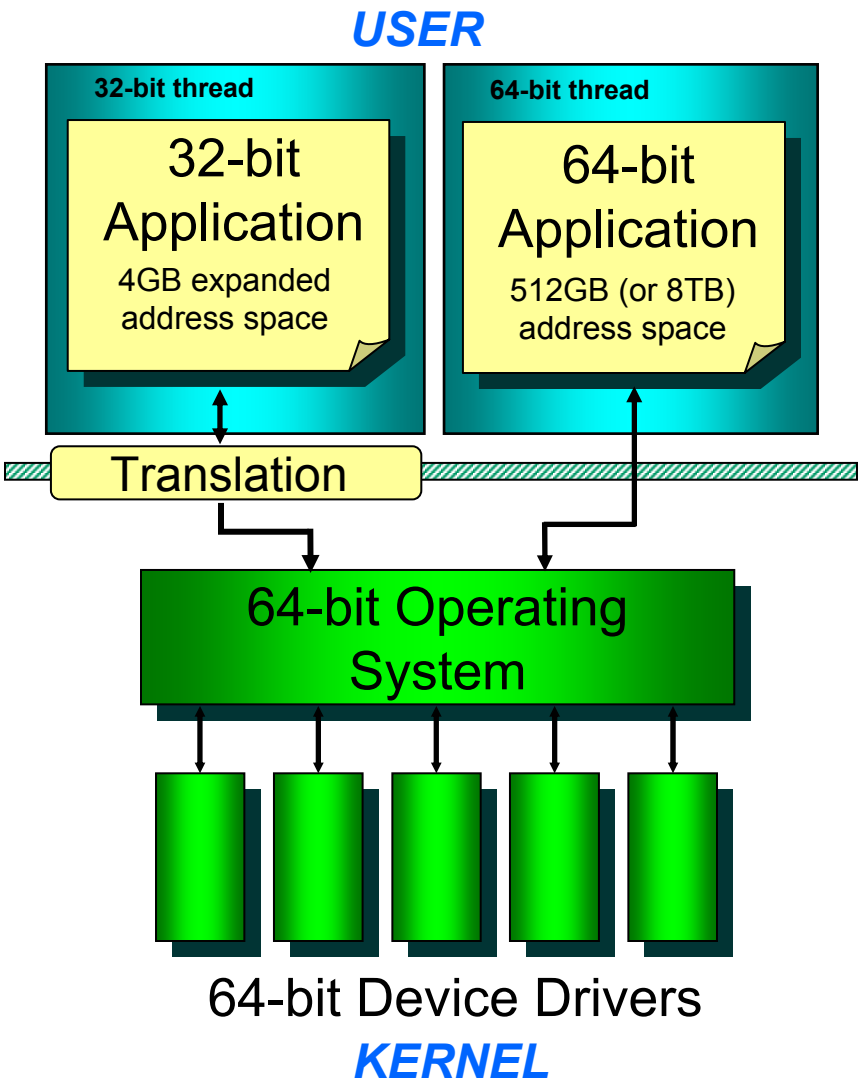
64-Bit Mode Operation

- Default data size is 32 bits
  - Override to 64 bits using new REX prefix
  - Override to 16 bits using legacy operation size prefix (66h)
- Default address size is 64 bits
  - Pointers are 64 bits
- 2 New instructions added, Some redundant encodings reclaimed
  - MOVSXD: Move sign extended double to quad
  - SWAPGS: Allows quick swap of GS in ISRs
- New override (REX) allows naming 16 GP and 16 SSE registers
  - Only 1 override byte per-instruction is needed for extended registers; regardless of how many are used by the instruction

64-bit OS & Application Interaction

32-bit Application Execution

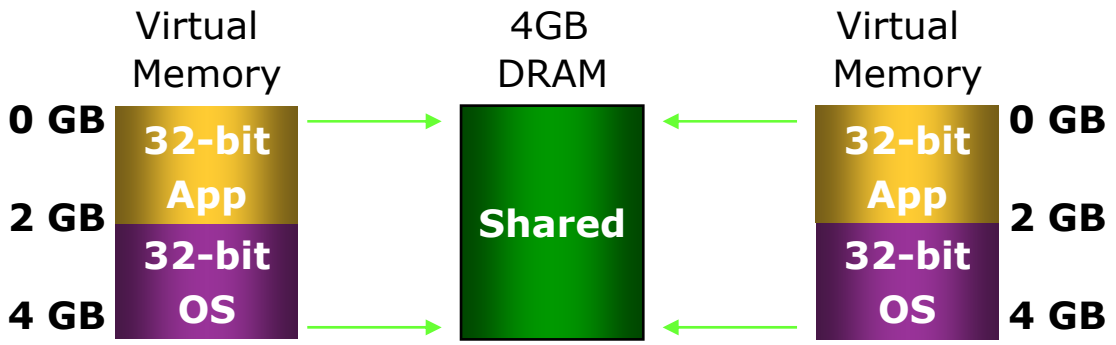
- No recompile required, 32-bit code directly executed by CPU
- 64-bit OS provides 32-bit libraries and "thunking" translation layer for 32-bit system calls.



Increased Memory for 32-bit Applications

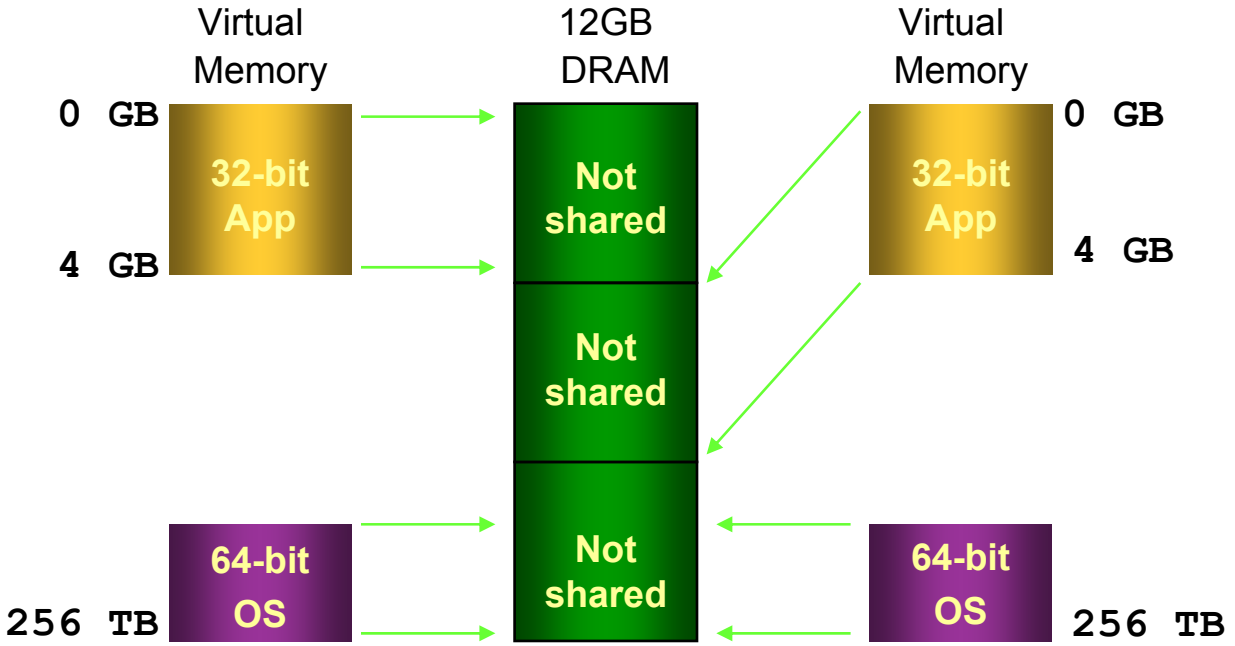
32-bit server, 4 GB DRAM

- OS & App share small 32-bit VM space
- 32-bit OS & applications all share 4GB DRAM
- Leads to small dataset sizes & lots of paging



64-bit server, 12 GB DRAM

- App has exclusive use of 32-bit VM space
- 64-bit OS can allocate each application large dedicated portions of 12GB DRAM
- OS uses VM space way above 32-bits
- Leads to larger dataset sizes & reduced paging



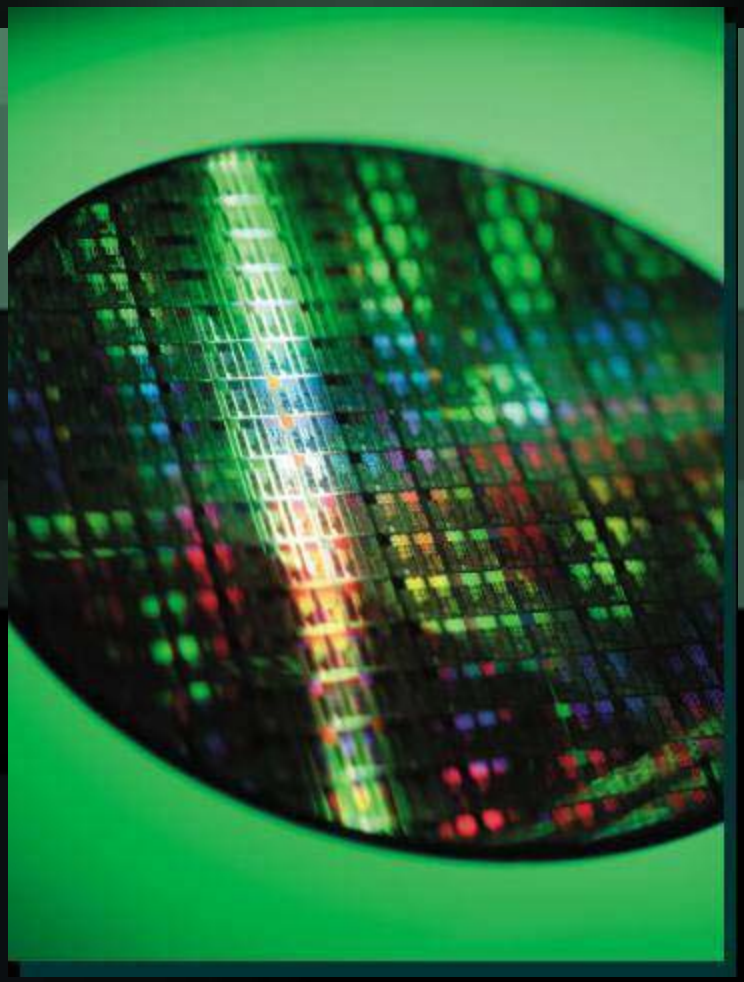
AMD Opteron™ Processor Ecosystem

32-bit software support: Not required to port!

- **Immediately compatible because AMD Opteron™ natively supports x86 instruction set**
- **Able to take full advantage of core enhancements offered by AMD Opteron™**
- **Gets increased performance today as seen by 3rd party benchmarks**

64-bit software support

- **Kernel-level programs must be ported, such as device drivers**
- **Programs that will share in-memory data structures with a 64-bit program must be ported**
- **Operating systems, development tools, and device drivers are ported, which provide the necessary infrastructure for running a mixture of 32/64 applications**
- **Applications that handle lots of data, users, and transactions**

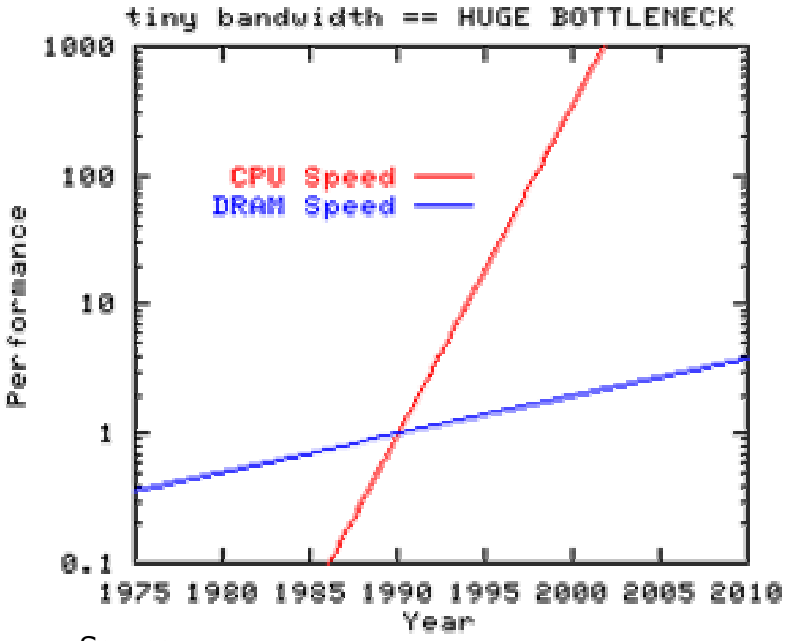


AMD64
Processor Design

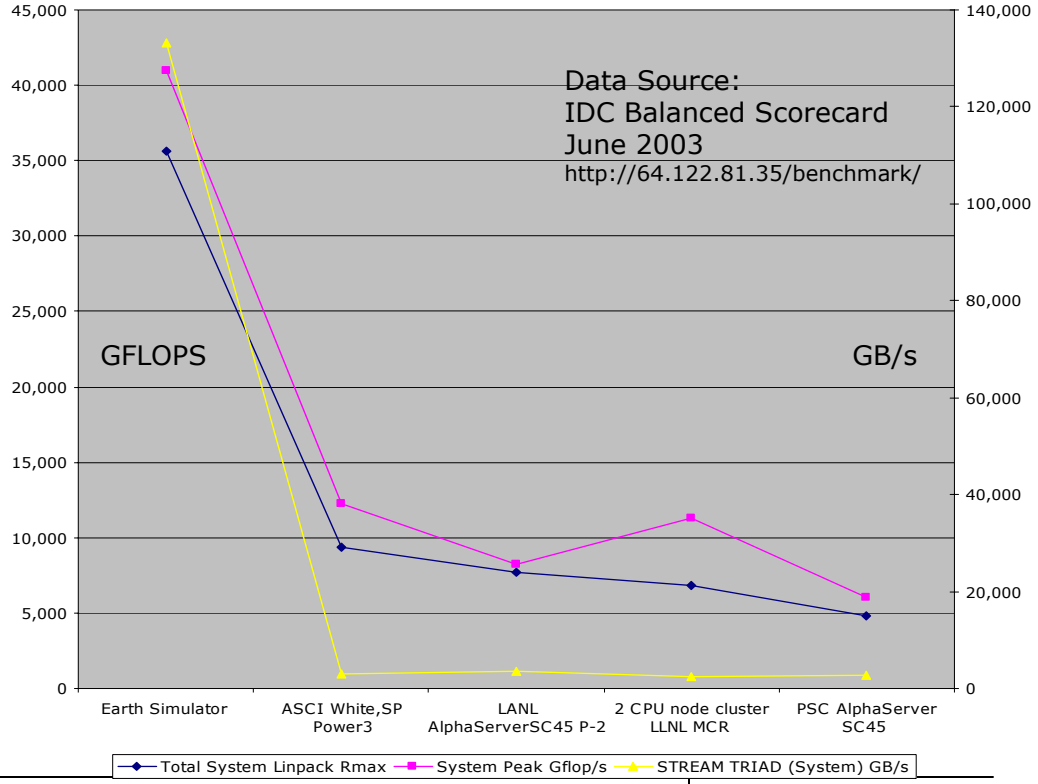
GFLOPS outpaced Bandwidth

Workshop on the Road Map for the Revitalization of High End Computing (June 18-23, 2003)

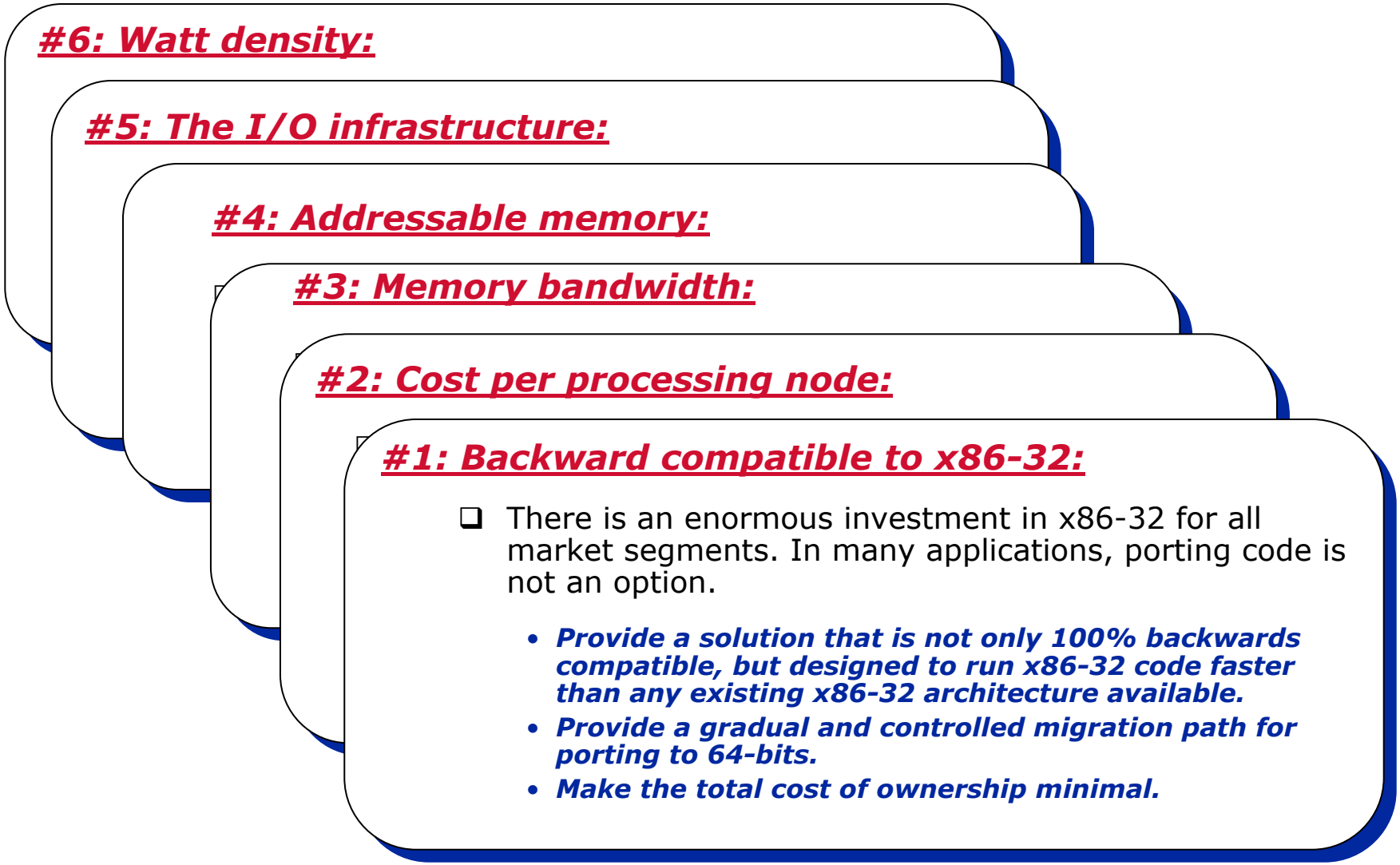
- Improve memory bandwidth
  - Sustained Memory bandwidth is not increasing fast enough
- Improve the Interconnects-"connecting to the interconnect"
  - Connecting through I/O is too slow we need to connect to CPU at memory equivalent speeds



Source:
John McCalpin STREAM
http://www.cs.virginia.edu/stream/stream_logo.gif



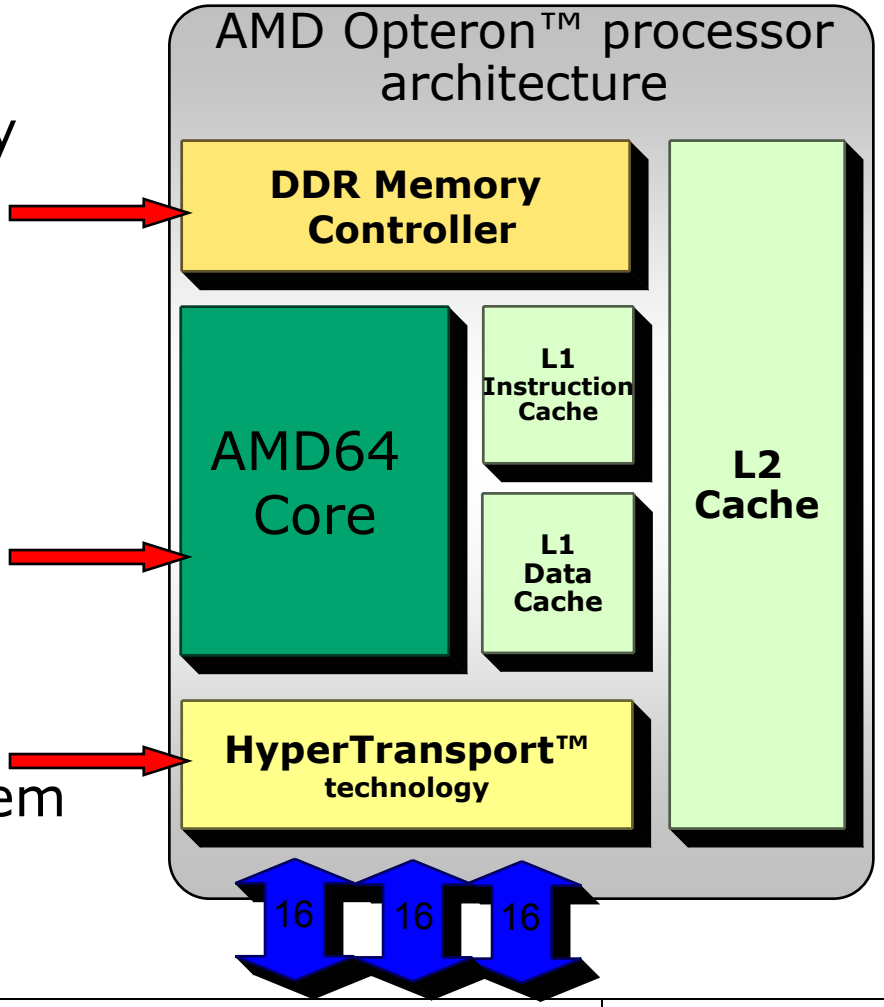
x86 in High Performance Computing
- *The Six System Challenges*



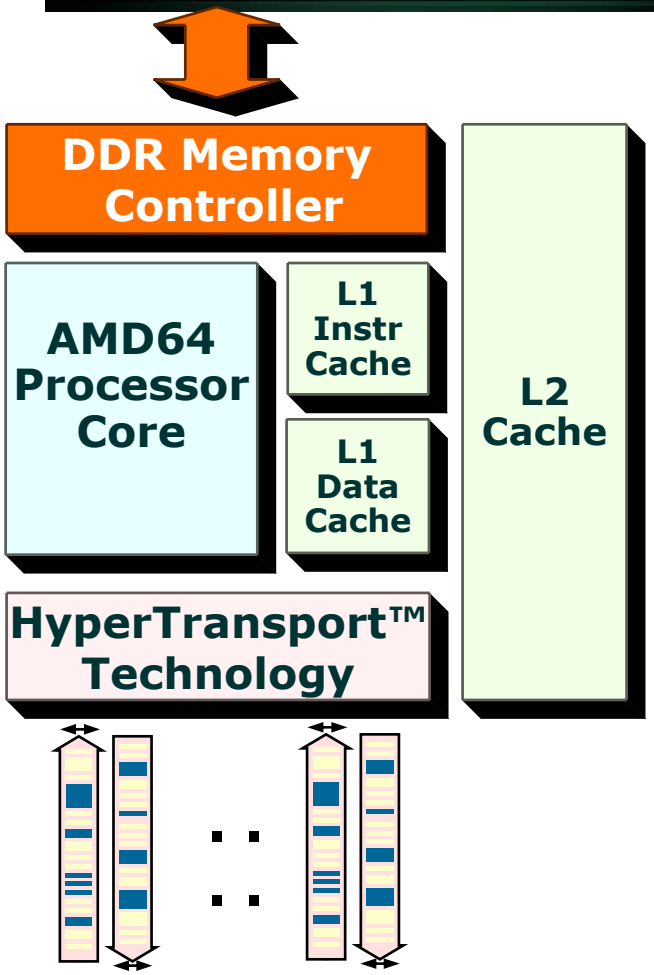
x86-64 Architecture Innovations

AMD Opteron™ performance-enhancing features:

- ***Performance*** - High-bandwidth integrated memory controller scales with processor frequency and number of processors
- ***Compatibility*** - Optimized 64-bit core with uncompromising support for 32-bit applications
- ***Scalability*** - Highly scalable HyperTransport™-based system bus enables glueless multiprocessing

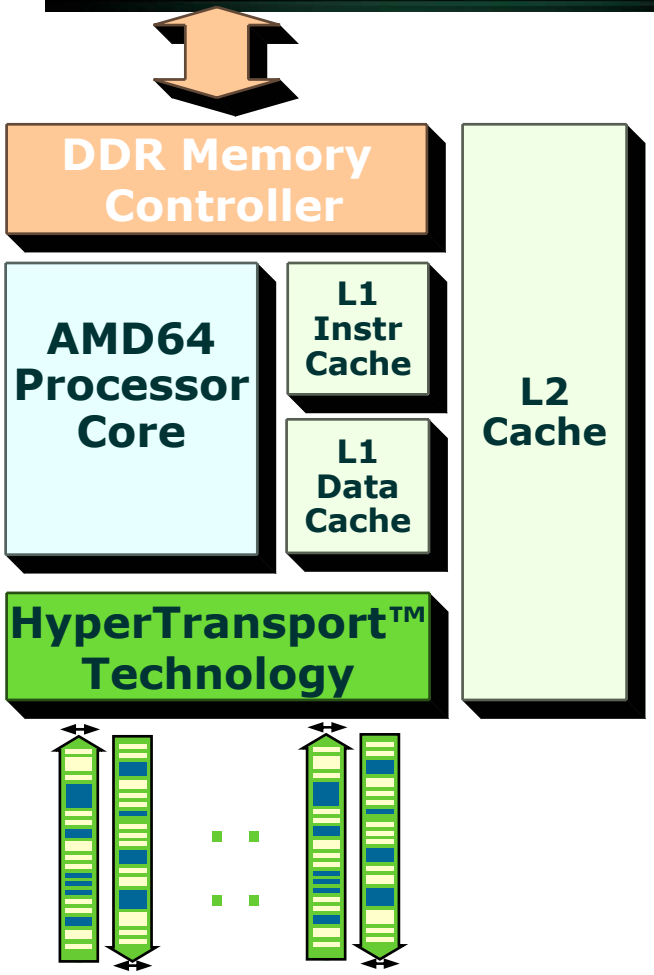


AMD64 Processors: Integrated Memory Controller



- **Run memory controller at processor speeds rather than FSB speeds**
  - **Today's AMD Athlon™ XP processor north bridge memory controllers run at 133 MHz**
- **Dramatically decrease latency**
  - **AMD64 architecture designed to achieve under 80 ns best case latency**
  - **Latency generally decreases further as the core frequency increases**
- **Add intelligence w/o decreasing performance**
- **Supports variety of DDR memories**
  - **200, 266 and 333 MHz**
  - **Registered and unbuffered DIMMs**
  - **Future processor cores planned to support DDR-II, etc.**

AMD64 Processors: HyperTransport™

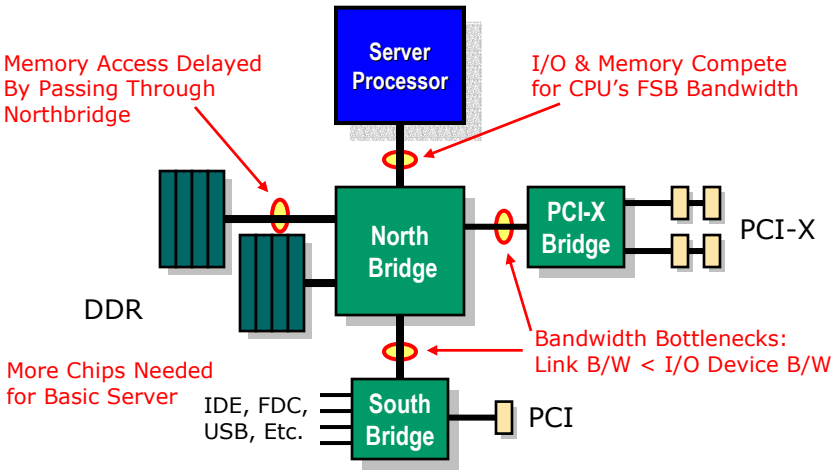


W = 2, 4, 8, 16, or 32-bits each way

- **HyperTransport™ technology is**
  - **High-speed, low pin-count, asynchronous, chip-to-chip board level interconnect**
  - **Proven technology in production today**
- **HyperTransport is not**
  - **A replacement for PCI, PCI-X, etc**
  - **A networking fabric**
- **HyperTransport physical interface**
  - **Point to point, differential, low-voltage swing**
  - **HyperTransport 1.0 -> Up to 1600MT/s to 12”**
  - **HyperTransport 2.0 -> Beyond 4000MT/s**
- **HyperTransport logical interface**
  - **100% PCI compliant API**
  - **OS I/O (PCI) enumeration code untouched for AMD64 processor based systems**

Advanced AMD Opteron™ Processor System Architecture

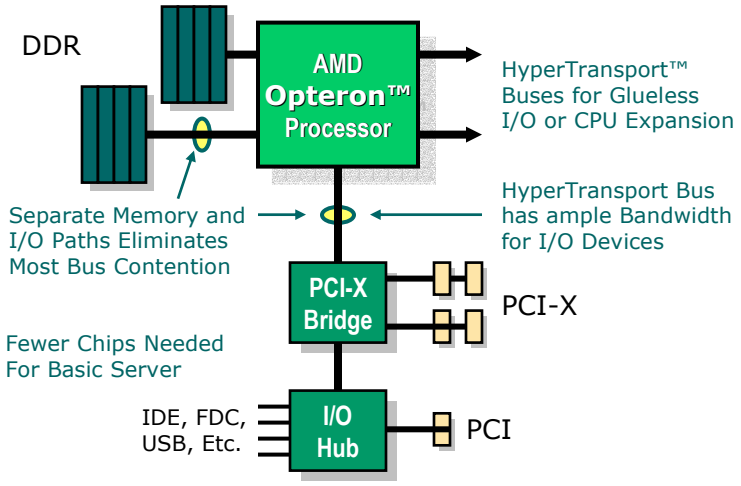
Typical System



- Must access memory through Northbridge
  - Longer latency memory access
- Memory and I/O access on the same bus
  - I/O and Memory compete for bandwidth
- Memory or I/O paths originate from Northbridge
  - Bandwidth does not scale well with more CPUs
- System logic uses more chips and many buses
  - Systems cost more to design, build and test

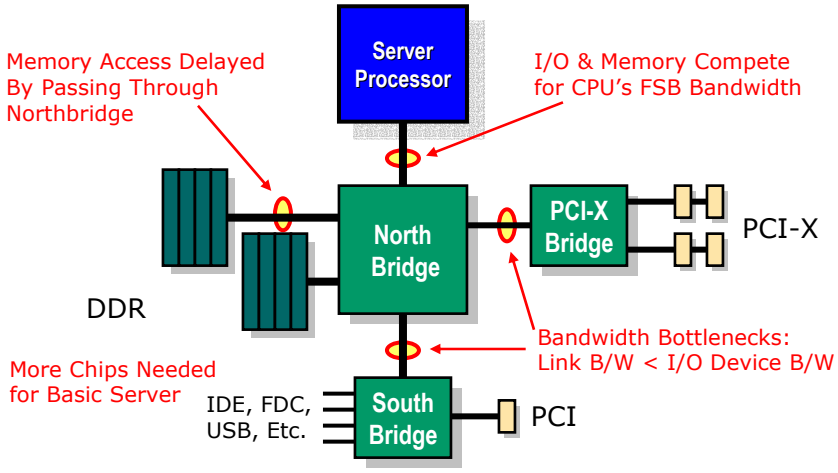
Advanced AMD Opteron™ Processor System Architecture

AMD64 System



- Integrated memory controller
  - Low latency memory access speeds processing
- Separate Memory and I/O pathways
  - Eliminates I/O and memory bus competition
- Each processor has more memory & I/O paths
  - Memory and I/O bandwidth scales well
- Modular glueless logic using HyperTransport™ bus
  - Fewer chips and lower cost implementation

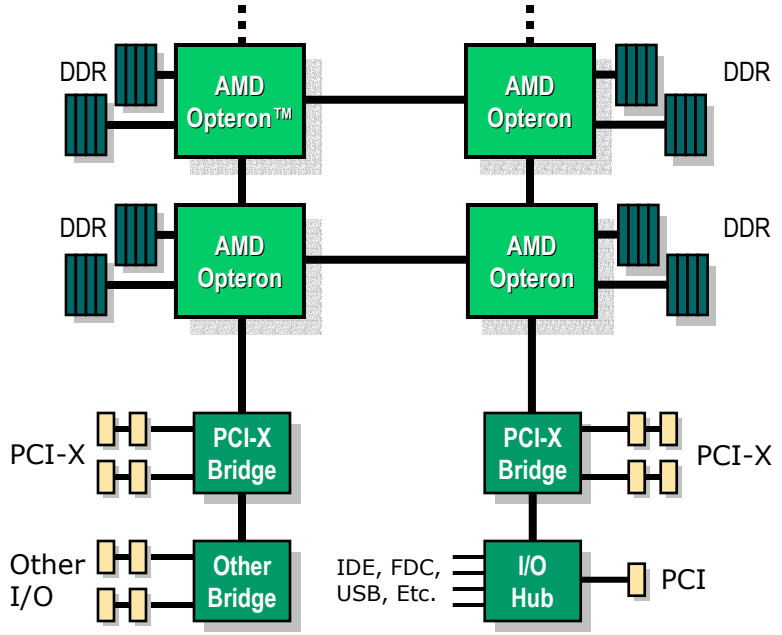
Typical System



- Must access memory through Northbridge
  - Longer latency memory access
- Memory and I/O access on the same bus
  - I/O and Memory compete for bandwidth
- Memory or I/O paths originate from Northbridge
  - Bandwidth does not scale well with more CPUs
- System logic uses more chips and many buses
  - Systems cost more to design, build and test

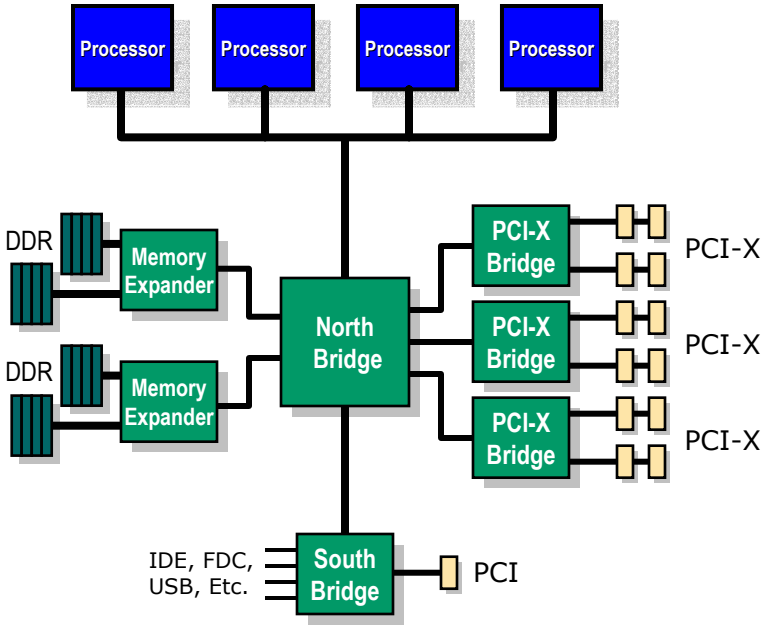
Advanced AMD Opteron™ Processor MP System Architecture

AMD Opteron™ System



- Scalable memory and I/O bandwidth
  - Up to 8 processors without glue logic
  - Each processor adds more memory
  - Each processor adds additional HyperTransport™ buses for more PCI-X and other I/O bridges
  - Fewer chips required

Typical MP System



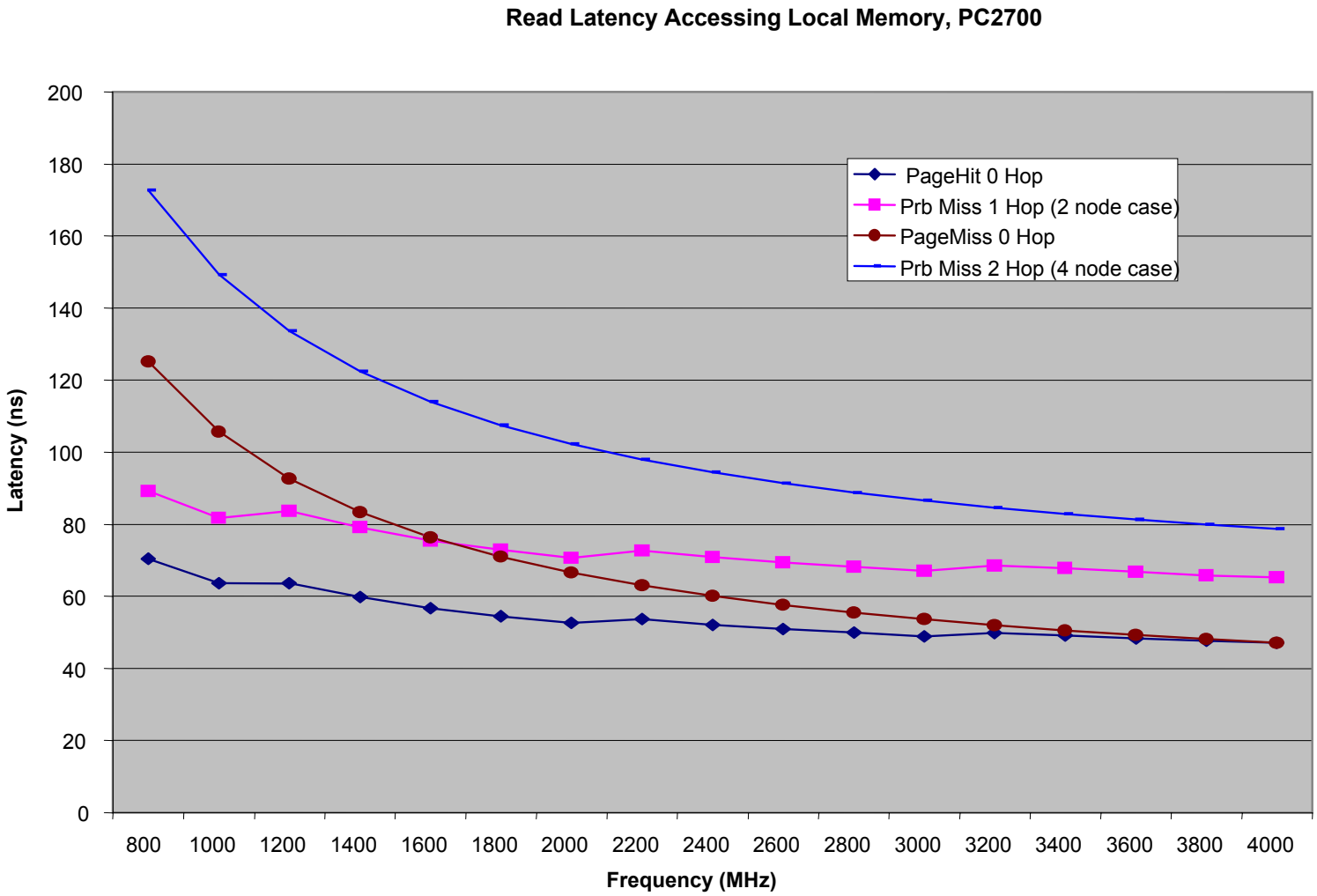
- System scalability limited by Northbridge
  - Maximum of 4 processors
    - Processors compete for FSB bandwidth
  - Memory size and bandwidth are limited
  - Maximum of 3 PCI-X bridges
  - Many more chips required

Simple SW Model for NUMA

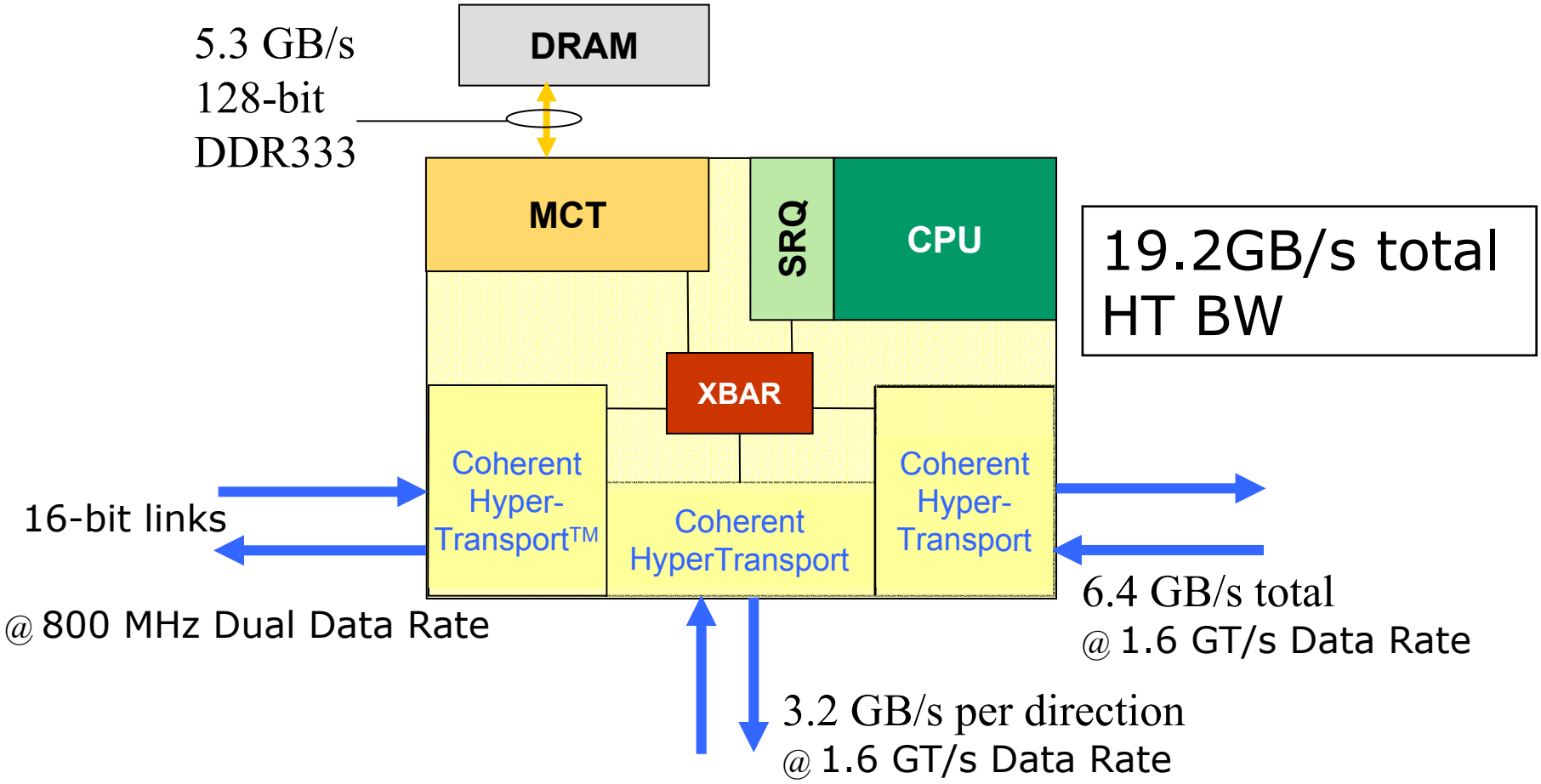
- NUMA brings scalability advantages
  - Scalable memory access, but requires thread affinity management
- SMP systems bring simplified software model
  - But does not scale without significant cost
- AMD Opteron™ processor provides benefits of both:
  SMP view for Software
  - Physical address space is flat and fully coherent
  - Latency difference between local and remote memory in an 8P system is comparable to the difference between a DRAM page hit and page miss
  - DRAM can be contiguous or interleaved
- MP support designed into processor & system from the beginning
  - Lower overall system chip count generally increases reliability and lowers cost
  - All MP system functions use CPU technology and frequency
- Latency shrinks quickly by increasing CPU and HyperTransport™ technology link speed
  - Additional processor nodes bring increased memory bandwidth and great overall system throughput

Integrated Memory Controller Latency

(Local memory access, registered memory, CAS2.5)



AMD Opteron™ Processor Architecture

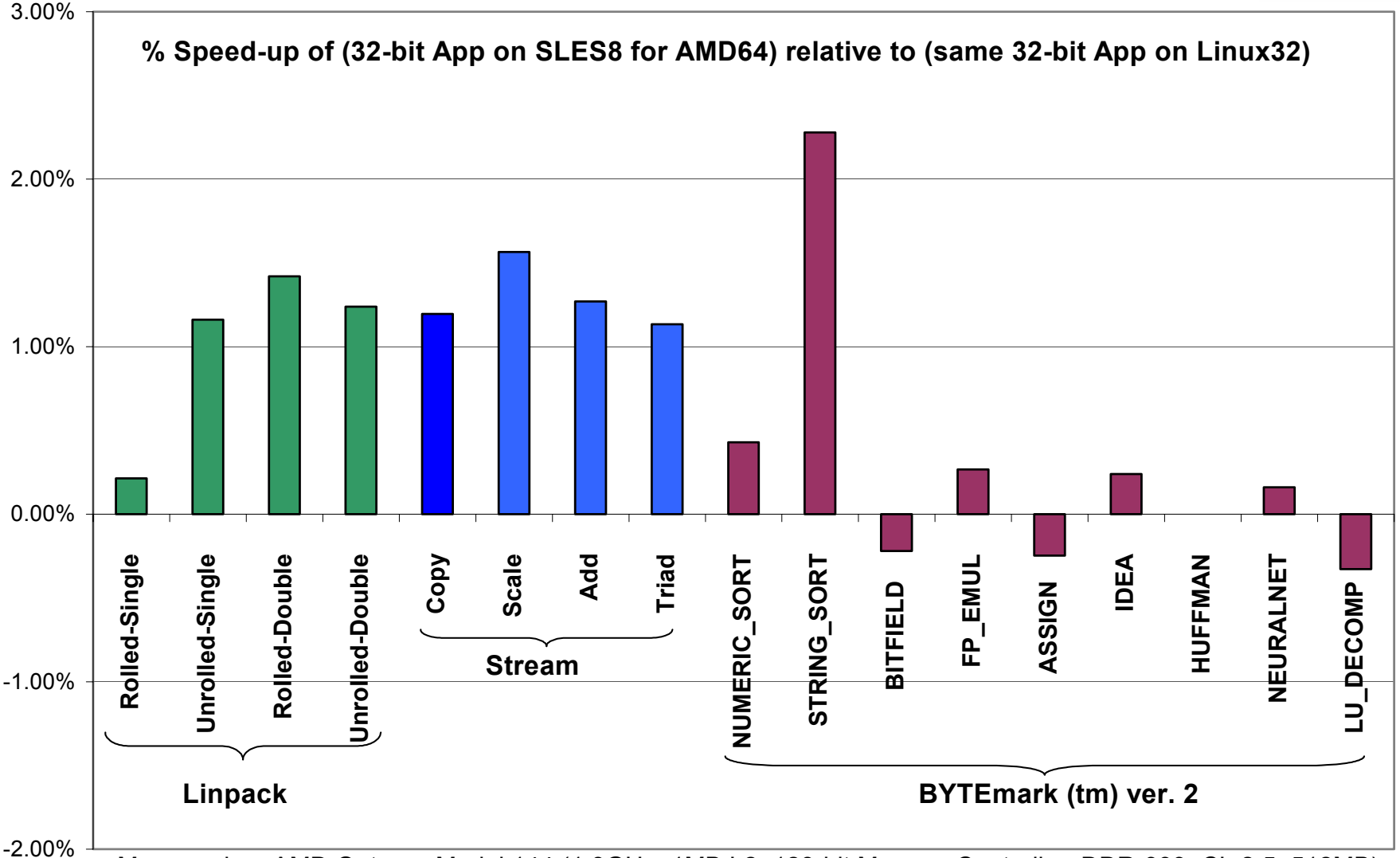


AMD64 Performance

32-bit OS & Application Support

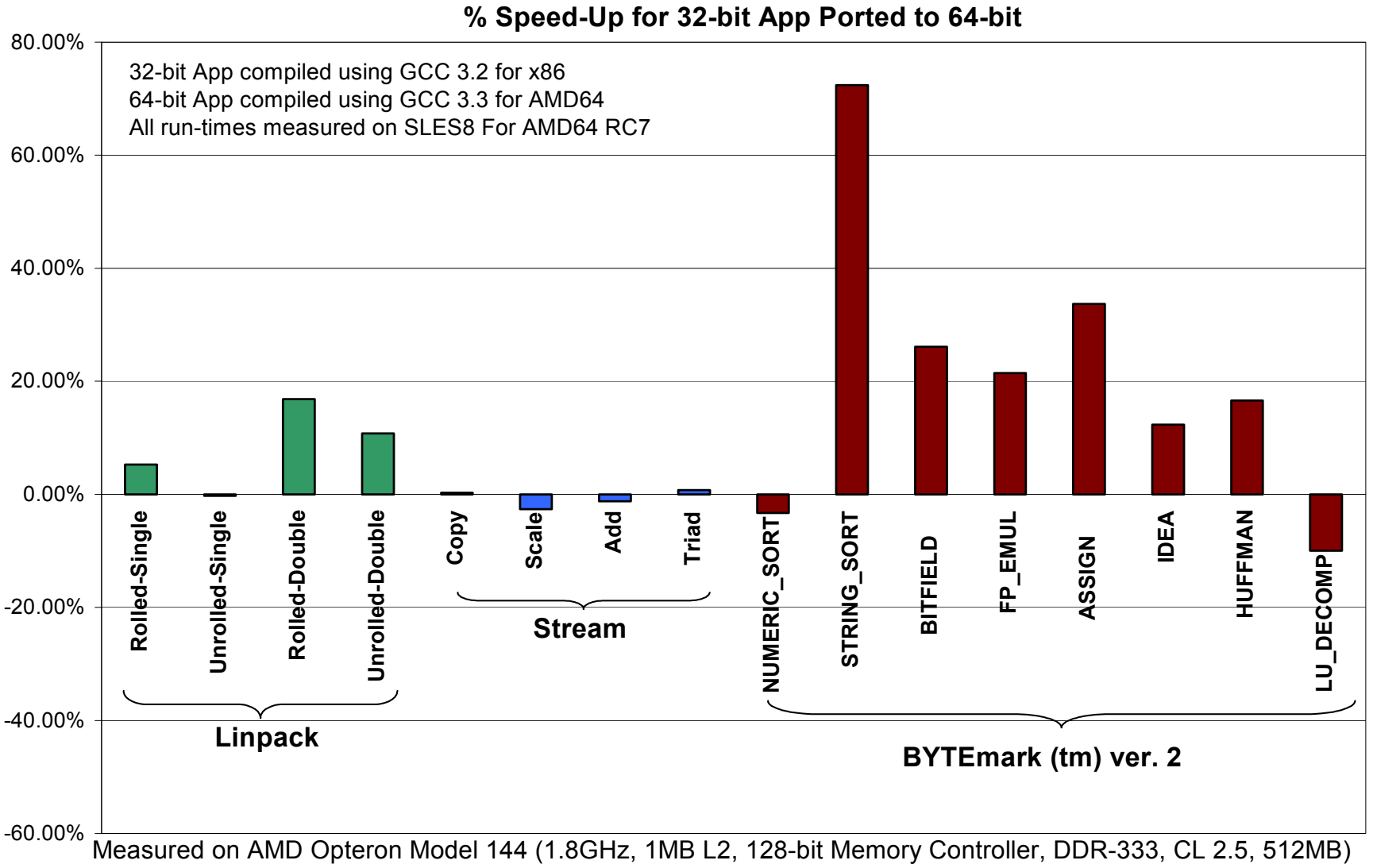
- 32-bit software does not have to be ported since the AMD Opteron™ processor is compatible with 32-bit OS, applications, and drivers
  - Natively supports x86 instruction set
  - Floating-point model is x87
  - SSE, SSE2, MMX™, 3DNow!™, and in-line ASM are supported
- AMD Opteron processor offers a high-performance platform for 32-bit software
  - Takes advantage of core enhancements offered by AMD Opteron processor
  - Is expected to progressively run faster as systems speed up
  - AMD Opteron processor-based systems are posting leading 32-bit benchmarks, including SPECint_rate, SPECfp®_rate, SPECWEB99, MMB2, and TPC-C

32-bit App Performance on 64-bit Linux



Measured on AMD Opteron Model 144 (1.8GHz, 1MB L2, 128-bit Memory Controller, DDR-333, CL 2.5, 512MB)

32-bit vs 64-bit App Performance



Measured on AMD Opteron Model 144 (1.8GHz, 1MB L2, 128-bit Memory Controller, DDR-333, CL 2.5, 512MB)

High Performance Linpack

GOTO Library Results AMD Opteron™ system	# P	Rmax (GFlops)	Nmax (order)	N1/2 (order)	Rpeak (GFlops)	GFLOP/ Proc	Rmax / Rpeak
4P AMD Opteron 1.8GHz 2GB/proc PC2700 8GB Total	**4**	**12.06**	**28000**	**1008**	**14.4**	**3.02**	**83.8%**
2P AMD Opteron 1.8GHz 2GB/proc PC2700 4GB Total	**2**	**6.22**	**20617**	**672**	**7.2**	**3.11**	**86.4%**
1P AMD Opteron 1.8GHz 2GB PC2700	**1**	**3.14**	**15400**	**336**	**3.6**	**3.14**	**87.1%**

High-Performance BLAS by Kazushige Goto

- sgemm/dgemm/cgemm/zgemm available today
- Optimized http://www.cs.utexas.edu/users/flame/goto

GOTO results were with 64-bit SuSE 8.1 Linux Professional Edition with NUMA kernel and Myrinet MPIch-gm-1.2.5..10 message passing library.

AMD Opteron™

Three Major Performance Innovations

AMD64 Technology

- Introducing 64-bit technology that is an extension of the x86 32-bit world
- No "Migration Tax"

Integrated Memory Controller

- No longer in the Northbridge, the memory controller is now in the processor
- Reduces memory performance bottlenecks

HyperTransport™ Interface

- Driving on-board chip-to-chip communication to new levels of performance
- Lower cost PCB solutions